SYSTEM AND METHOD FOR SPEAKERPHONE OPERATION IN A COMMUNICATIONS DEVICE

FIELD OF THE INVENTION

[0001]   The invention relates to the field of communications, and more particularly to techniques for generating clearer and more reliable speakerphone operation in a cellular telephone or other communications device.

BACKGROUND OF THE INVENTION

[0002]   Convenient and effective speakerphone operation has become a desirable feature in cellular handsets and other communications devices.   Communities concerned with traffic safety have in some instances banned the handheld operation of cellular phones while driving.   Handsets and other devices equipped with a speakerphone feature permit users to place the device in a resting position in a car or other location while still carrying out normal conversations and other telephone access.

[0003]   However, equipping a cellular telephone with an effective speakerphone capability is not a trivial integration task.   One practical difficulty is that many cellular telephones are small devices which contain both an earpiece speaker and integrated microphone within a few inches of each other, to make the unit more compact. Therefore, duplex-type operation where both the speaker path and microphone path are active at the same time may generate unwanted feedback, since the output of the speaker leaks into the microphone via air and case vibration. This feedback problem only gets worse as speaker volumes are increased, such as they might be in a noisy car or room.

[0004]   Echo canceling circuits are known which can be connected to the microphone path on a cellular phone or other device, and remove a portion of the feedback energy emanating from the speaker. Unfortunately, echo canceling circuits are currently only capable of about 35 dB of cancellation, and the energy from the speaker may be more than 35 dB greater than the energy delivered by the embedded microphone so that echo and feedback still occur, even when echo cancellation circuits are included.

[0005]   One solution to the speakerphone problem is to attempt to physically isolate the speaker and microphone from each other in the handset. For instance, one may place the speaker used for speakerphone operation in a rear-facing part of the handset so that less sound impinges directly on the microphone from the speaker. However, this placement makes the sound harder to hear for a user from whom the speaker faces away, and some amount of speaker energy will still leak through the cellular or other case to the microphone.

[0006]   Another solution to feedback is to prevent the speaker path and microphone path from operating at the same time. This simplex-type of operation makes direct feedback impossible but results in one-way communication only, which requires users at both ends to signal the end of their speech, and wait for a response. More effective and natural speakerphone operation is desirable. Other problems exist.

SUMMARY OF THE INVENTION

[0007]   The invention overcoming these and other problems in the art relates in one regard to a system and method for speakerphone operation in a communications device, in which built-in intelligence simultaneously manages both the speaker path and the microphone path of the device to reduce unwanted echo and feedback while

2

still preserving a perceived quality of conversational speech. In an embodiment of the invention, a communications device such as a cellular telephone handset or other device may incorporate dual voice activity detection circuits to simultaneously monitor the signal energy and other characteristics in both speaker and microphone paths, and award control to one or the other path based on dynamic thresholds or other adaptive or other criteria. In other embodiments, problems such as premature dropouts caused by greater than average background noise may be prevented by applying hangtime parameters which keep the speaker path open until a minimum interval has passed, before transferring control to the microphone path. The criteria applied to trigger a change in control from speaker path to microphone path and vice versa may also be adapted in embodiments of the invention, including to eliminate a lower threshold below which the speaker path switches out and passes control to the microphone path, automatically.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008]   The invention will be described with reference to the accompanying drawings, in which like elements are referenced with like numbers, and in which:

[0009]   Fig. 1 illustrates a two-way communications platform including speakerphone operation, according to an embodiment of the invention.

[00010]   Figs. 2(A)-2(C) illustrate processing of inbound and outbound speech in different regards, according to an embodiment of the invention.

[00011]   Fig. 3 illustrates a speakerphone control operation, according to an embodiment of the invention.

86768v1

[00012]   Figs. 4(A) and 4(B) illustrate processing of inbound and outbound speech in different regards, according to an embodiment of the invention.

[00013]   Fig. 5 illustrates inbound and outbound speech envelopes, according to an embodiment of the invention.

[00014]   Fig. 6 illustrates a dynamic inbound break-in threshold and other speech processing, according to an embodiment of the invention.

[00015]   Fig. 7 illustrates inbound break-in instances using a dynamic break-in threshold and other speech processing, according to an embodiment of the invention.

[00016]   Fig. 8 illustrates a speakerphone control operation, according to an embodiment of the invention.

[00017]   Figs. 9(A) and 9(B) illustrate processing of inbound and outbound speech in different regards, according to an embodiment of the invention.

[00018]   Figs. 10(A) and 10(B) illustrate outbound and inbound path control including an interposed hangtime, according to an embodiment of the invention.

[00019]   Fig. 11 illustrates a speakerphone control operation, according to an embodiment of the invention.

[00020]   Figs. 12(A) and 12(B) illustrate processing of inbound and outbound speech in different regards, according to an embodiment of the invention.

[00021]   Fig. 13 illustrates speaker path activation, according to conventional far-end processing during noisy conditions.

[00022]   Figs. 14(A) and 14(B) illustrate speaker path activation during noisy conditions, according to an embodiment of the invention.

DETAILED DESCRIPTION OF EMBODIMENTS

[00023]   Fig. 1 illustrates an architecture of a communications device having a speakerphone capability according to an embodiment of the invention. The device illustrated in Fig. 1 may be or include, for instance, a cellular telephone handset, a voice-enabled wired or wireless device such as a networked Voice over IP (VoIP) or ISDN telephone device, a two-way radio communications device, a modem or hybrid telephone/modem device, a wired or wireless telephone connected to the public switched telephone network (PSTN) via a speakerphone base, or other communications devices or platforms. In general, according to the illustrated architecture the communications device may include a microphone path 128 which includes a microphone 102 or other acoustical or other input transducer, and a speaker path 130 which includes a speaker 120 or other acoustical or other output transducer. In embodiments, in general only one of the microphone path 128 and the speaker path 130 may be activated at the same time, to avoid feedback between the two transducers. Other modes are possible in other embodiments. The microphone path 128 may from time to time be referred to as the inbound or near-end channel, and the speaker path 130 as the outbound or far-end channel, respectively.

[00024]   The microphone 102 in the microphone path 128 may be connected to a microphone gain control 104, to boost or attenuate the output of microphone 102 as appropriate. The output of the microphone gain control 104 may be communicated to an echo canceller 106 to remove a portion of any feedback, including echo, leaking from speaker 120 to microphone 102. Echo canceller 106 may for example be implemented in hardware, software, firmware of a combination thereof. Echo canceller 106 may for instance be implemented instance using commercially available

4

parts such as dedicated integrated circuits manufactured by Oki Semiconductor or others, or using software modules such as echo canceller modules available for digital signal processors such as the DSP 56000 family manufactured by Motorola Corp., digital signal processors made by Texas Instruments Inc., or others. In embodiments, the echo canceller 106 may incorporate or implement known echo cancellation algorithms, for instance algorithms related to or incorporated in International Telecommunications Union (ITU) standard G.165 or other cancellation algorithms or techniques. In embodiments, the echo canceller 106 may reduce the echo or other feedback by as much as 35 dB or more, but may typically not eliminate the full degree of feedback present in the signal generated by the microphone 102.

[00025]   The output of the echo canceller 106 may be communicated to a speech encoder 108, which compresses or otherwise processes speech input for purposes of wireless or other transmission. The speech encoder 108 may be implemented using known speech compression or other algorithms, for instance algorithms related to or incorporated in ITU standards such as ITU G.711, G.723, G.726, G.729, or other protocols. Those standards or protocols may incorporate or implement for example the Low-Delay Code-Excited Linear Prediction (LD-CELP) speech coding algorithm, which encodes 2.5 ms frames of digitized, telephone bandwidth speech or audio signals sampled at 8 KHz, or other digitizing or other techniques. Other speech compression/decompression (codec) algorithms, software or standards may be used. The speech encoder 108 may likewise be implemented in hardware, software, firmware or a combination thereof, including using programmable digital signal processors or other components.

5

[00026]   After a user's speech input is encoded by the speech encoder 108, the encoded speech may be communicated to the modem transmit module 110. The modem transmit module 110 may prepare the encoded signal for wireless or other transmission via an antenna or other air or other interface, for instance generating wireless transmission in the 800/900 MHz, 1.9GHz or other cellular, PCS or other frequency spectra for voice or other communications.

[00027]   On the receiver side, a modem receiver module 126 may likewise be coupled to a cellular antenna or other source of radio frequency (RF) or other wireless or other energy to capture, downconvert and/or demodulate wireless carrier signals. The modem receive module 126 may communicate the demodulated received signal to a speech decoder 124. The speech decoder 124 may in general perform the reverse type of operation from the speech encoder 108, for example to decompress far-end speech from a remote user of another cellular handset or other device. The output of speech decoder 124 may be communicated to the speaker gain control 122, providing amplification or attenuation of the decoded speech for driving the speaker 120, such as the earpiece speaker in a cellular handset or other transducer. The output of the speech decoder 124 may also be communicated to the echo canceller 106 to perform echo detection and cancellation processing.

[00028]   In embodiments of the invention such as that illustrated in Fig. 1, the microphone path 128 and the speaker path 130 may each be coupled to further circuitry to monitor and manage the speakerphone operation of the communications device. More specifically, the output of the echo canceller 106 may also be communicated to an inbound voice activity detector (VAD) 114. The output of the speech decoder 124 may similarly be communicated to an outbound voice activity

6

detector (VAD) 118. Each of inbound VAD 114 and outbound VAD 118 may also be implemented using hardware, software, firmware of a combination thereof. The inbound VAD 114 and outbound VAD 118 may, for instance, each be implemented using a microprocessor, a digital signal processor or other processors. The VAD 114 and VAD 118 may each generate a speech energy envelope, speech sample, voice-present or other types of speech detection signals or functions used to identify the presence of speech information, as opposed to background or other types of noise. Inbound VAD 114 and outbound VAD 118 may for instance be programmed to perform speech detection algorithms, such as those related to or incorporated in ITU standards or others, for instance according or related to the ITU G.711, G.723, G.726, G.729 or other standards. The inbound VAD 114 and outbound VAD 118 may also be coupled together, to permit direct communication therebetween.

[00029]   The output of each of the inbound VAD 114 and the outbound VAD 118 may in turn be communicated to a duplex arbiter 116. Duplex arbiter 116 may also be implemented using hardware such as a microprocessor or digital signal processor, in software, firmware or a combination thereof to perform supervisory tasks to arbitrate and manage the activation of the microphone path 128, speaker path 130 and other resources to enhance speakerphone and other operation. The duplex arbiter 116 may, for instance, determine instances in time when the inbound (near-end, or handheld user of the communications device) speech energy is significant while the outbound (far-end, or remote user) speech energy is negligible so that the duplex arbiter 116 may activate the microphone path 128 to capture that local speech, while deactivating or muting the speaker path 130 since the far-end user is interpreted as not speaking or communicating.

7

[00030]   Conversely, in instances when the inbound speech energy detected by the inbound VAD 114 is negligible while the outbound speech energy detected by the outbound VAD 118 is significant, the duplex arbiter 116 may activate the speaker path 130 while deactivating the microphone path 128, so that the far-end user's speech may be heard over the speaker 120.

[00031]   On the other hand, during those intervals of time in which both the inbound VAD 114 and outbound VAD 118 detect significant speech energy in their respective paths, the duplex arbiter 116 may apply selective criteria to decide which path to activate. As illustrated for instance in Figs. 2(A) – 2(C), intervals may occur when both the inbound VAD 114 (Fig. 2(B)) and outbound VAD 118 (Fig. 2(A)) have detected speech energy greater than their respective detection thresholds, and present duplex arbiter 116 with a speech-detected signal, illustrated as a gate function.

[00032]   As illustrated in Fig. 2(C), when both VAD signals are active, the duplex arbiter 116 may choose to activate one or the other path. As illustrated in that figure, in embodiments the duplex arbiter 116 may switch control to the microphone path 128 (inbound channel) when speech is recognized at the microphone 102, even when the absolute value of the energy presented by the presumed speech signal is less than the output of the outbound VAD 118. This decision criteria may be applied because the energy of the speech content in the microphone path 128 may typically be significantly less than that of the speaker path 130, even when a user is speaking with a normal voice close to the microphone 102, which intensity only decreases when the cellular handset or other device is placed farther away from the user.

[00033]   Operation of this type may permit seamless transitions between the near-end and far-end user's speech in conversation, and prevent artifacts such as channel

8

lockouts. In embodiments, as illustrated the duplex arbiter 116 may also communicate with a comfort noise generation and substitution module 112, likewise capable of being implemented in hardware, software or firmware or a combination thereof. The comfort noise generation and substitution module 112 may in turn also communicate with the microphone gain control 104 and the speaker gain control 122, to output white noise or other comparatively pleasant or innocuous sounds during path transitions, dead spots such as when both the microphone path 128 and speaker path 130 may be muted, or at other times. In other embodiments or under other conditions, the duplex arbiter 116 may award control to the microphone path 128 or the speaker path 130 under different fixed or dynamic criteria used for decision processing.

[00034]   In an embodiment illustrated in Fig. 3, for example, a threshold used to award control to the microphone path 128 may be dynamically computed based on the energy being produced by speech encoder and other parameters. In step 302, processing may begin. In step 304, microphone samples from the microphone 102 and speaker samples from the speaker 120 may be communicated to the echo canceller 106. In step 306, the speech encoder 108 may process the output of echo canceller 106. In step 308, a break-in threshold, referred to as "ib_break_in_thresh" and used for deciding to award control to the microphone path 128 while muting the speaker path 130, may be dynamically computed based on the outbound speech (or speaker) energy for the present discrete speech frame (n) and speech encoder parameters. In embodiments, that calculation may be or include the following computations:

<u>Algorithm 1</u>

$ib\_break\_in\_thresh(n) = \beta*ob\_r0(n);$

9

IF *(ib_break_in_thresh(n) > ib_break_in_thresh(n-1))*
   ib_break_in_thresh(n) = β*ob_r0(n);
ELSE
   ib_break_in_thresh(n) = α*ib_break_in_thresh(n-1) + (1-
α)*β*ob_r0(n);
END

Where: ob_r0(n) = outbound speech energy for a frame n;
   n = current speech frame
   β = an energy scalar; and
   α = decay rate.

[00035] In step 310, the output of the speech encoder 108 may also be communicated to an inbound speech envelope generator 132, which may in embodiments be integrated with or interface to inbound VAD 114. Inbound speech envelope generator 132 may generate a moving envelope representing speech energy, such as a moving average or other representation of speech energy of the signal in the microphone path 128. Outbound speech envelope generator 134, which also may be integrated with or interface to outbound VAD 118, may similarly generate an envelope output based on the signal in the speaker path 130.

[00036] In step 312, the resulting speech envelope may be compared to the current inbound break-in threshold (ib_break_in_thresh). If the envelope of the inbound speech exceeds that threshold, processing proceeds to step 314 where the duplex arbiter 116 may mute the speaker path 130 and activate or unmute the microphone path 128, thus allowing the near-end user's speech to be captured and communicated to the far-end user. If the envelope of the inbound speech does not exceed the inbound break-in threshold (ib_break_in_thresh), processing proceeds to step 316 where processing for the current frame of time may end, following which processing may repeat, proceed to other tasks or end.

[00037]   Figs. 4(A) and 4(B) illustrate speaker samples and echo-cancelled microphone samples, respectively, generated according to the embodiment illustrated in Fig. 3. Fig. 5 depicts an illustrative speech envelope for the inbound and outbound signals generated according to that embodiment. As illustrated in that figure, at certain times the inbound signal may exceed the outbound signal, while at other times the outbound signal may be greater than the inbound signal.

[00038]   Fig. 6 illustrates an overlay of the outbound (speaker path 130) speech energy on an illustrative inbound dynamic break-in threshold, with a fixed inbound break-in threshold also shown for comparison. As illustrated in that figure, the inbound break-in threshold may be made a dynamic function of the parameters of Algorithm 1 or otherwise, resulting in a time-varying threshold which tracks, at least in part, the outbound speech energy with which the inbound speech is in competition. Thus, in intervals during which the outbound speech energy is comparatively high, the inbound break-in threshold rises to a relatively higher plateau, forcing near-end speech at the microphone 102 to be greater in intensity to capture the channel. Conversely, the inbound break-in threshold may be relaxed in intervals during which the outbound speech energy decreases, so that comparatively softer near-end speech may activate the microphone path 128, unlike the fixed threshold approach.

[00039]   Fig. 7 illustrates the inbound speech envelope, inbound break-in dynamic threshold and inbound break-in instances generated according to the embodiment shown in Fig. 3. As illustrated in that figure, the inbound break-in instances may consequently occur in those periods of time where a relatively quiet outbound channel has driven the inbound break-in threshold to a lower level, enabling the microphone path 128 to appropriately seize the channel even with less energetic speech.

11

[00040]   When encoded speech is choppy or contains large swings in amplitude or other artifacts, in cases those inputs may cause rapid switching between microphone path 128 and speaker path 130, or other "race" or other undesirable conditions. In an embodiment of the invention illustrated in Fig. 8, the duplex arbiter 116 and other cooperating components may insert a delay interval or hangtime before permitting a transition of control from the microphone path 128 to the speaker path 130, and vice versa. The introduction of a hangtime may serve to prevent such race conditions when one or both of the near-end and far-end speech contains rapidly varying amplitudes.

[00041]   As shown in Fig. 8, in step 802 processing may begin. In step 804, near-end samples from the microphone 102 may be processed by the speech encoder 108. In step 806, outbound speech from the far-end user may be processed by speech decoder 124. In step 808, the echo canceller 106 may receive the outputs of the speech encoder 108 and the speech decoder 124 to suppress echo and other feedback artifacts. In step 810, the echo-cancelled inbound speech and the decoded outbound speech may be communicated to inbound speech envelope generator 132 and outbound speech envelope generator 134, respectively, to generate speech energy envelopes or other functions.

[00042]   In step 812, an inbound break-in threshold (ib_break_in_threshold) and outbound break-in threshold (ob_break_in_threshold) may be generated, for instance according to the embodiment illustrated in Fig. 3 or otherwise. In step 814, at least one of an inbound hangtime (ib_hang_time) and an outbound hangtime (ob_hang_time) may be decremented, or set to initial values if the communications device is in an initialization mode such as in a startup or reset operation. In step 816,

a determination may be made whether the speaker path 130 is activated. If the speaker path 130 is not activated, processing may proceed to step 818 where a determination may be made whether the microphone path 128 is activated.

[00043]   If the microphone path 128 is not activated, processing may proceed to step 822 where the microphone path 128 may be activated or unmuted, while the speaker path 130 may be deactivated or muted. After step 822, control may proceed to step 840 where processing for the current frame may end, following which processing may repeat, proceed to other tasks or end.

[00044]   If the determination at step 818 is that the microphone path 128 is on, processing may proceed to step 820 where a determination may be made the outbound speech envelope (ob_env) may be greater than the outbound break-in threshold (ob_break_in_threshold). If the outbound speech envelope (ob_env) is greater than the outbound break-in threshold (ob_break_in_threshold), processing may proceed to step 824 where a determination may be made whether the inbound hangtime (ib_hang_time) has expired. If the inbound hangtime (ib_hang_time) has not expired, processing may proceed to step 822 where again the microphone path 128 may be activated or unmuted, while the speaker path 130 may be deactivated or muted.

[00045]   If at step 824 the inbound hangtime (ib_hangtime) has expired, processing may proceed to step 826 where an outbound hangtime (ob_hangtime) may be set to begin a hangtime period for the speaker path 130. The outbound hangtime (ob_hangtime) may for instance be set to a fixed amount of time, such as 4 seconds or another value according to implementation. In embodiments, the outbound hangtime may be computed or set on a dynamic basis, for instance as a function of prior inbound or outbound hangtimes, detected speech energy in the inbound or outbound

13

paths or other variables. In step 828, the microphone path 128 may be deactivated or muted, while the speaker path 130 may be activated or unmuted, after which control may proceed to step 840 where processing for the current frame of time may end, following which processing may repeat, proceed to other tasks or end.

[00046]   If at step 820 the outbound speech envelope (ob_env) is determined to not exceed the outbound break-in threshold (ob_break_in_threshold), processing may proceed to step 822 where again the microphone path 128 may be activated or unmuted, while the speaker path 130 may be deactivated or muted. Control may then also proceed to step 840 where processing for the current frame of time may end, following which processing may repeat, proceed to other tasks or end.

[00047]   If at step 816 a determination is made that the speaker path 130 is on, processing may proceed to step 830 in which a determination may be made whether the inbound envelope (ib_envelope) exceeds the inbound break-in threshold (ib_break_in_threshold). If the inbound envelope (ib_envelope) does not exceed the inbound break-in threshold (ib_break_in_threshold), processing may proceed to step 832 where the speaker path 130 may be activated or unmuted while the microphone path 128 may be deactivated or muted. Following that step, control may then proceed to step 840 where processing for the current frame of time may end, following which processing may repeat, proceed to other tasks or end.

[00048]   If at step 830 a determination is made that the inbound envelope (ib_envelope) exceeds the inbound break-in threshold (ib_break_in_threshold), processing may proceed to step 834 where a determination may be made whether the outbound hangtime (ob_hangtime) has expired. If the outbound hangtime (ob_hangtime) has not expired, processing may likewise proceed to step 832 where the speaker path 130

14

may be activated or unmuted while the microphone path 128 may be deactivated or muted.

[00049]   If at step 834 a determination is made that the outbound hangtime (ob_hangtime) has expired, processing may proceed to step 836 where the inbound hangtime may be set to a fixed amount of time, such as 4 seconds or another value according to implementation.   In embodiments, the inbound hangtime may be computed or set on a dynamic basis, for instance as a function of prior inbound or outbound hangtimes, detected speech energy in the inbound or outbound paths or other variables. Processing may then proceed to step 838, where the speaker path 130 may be deactivated or muted while the microphone path 128 may be activated or unmuted. Following that step, control may then proceed to step 840 where processing for the current frame of time may end, following which processing may repeat, proceed to other tasks or end.

[00050]   In the embodiment of the invention illustrated in Fig. 8, the awarding of control to the microphone path 128 or the speaker path 130 may therefore depend on more than one criterion. Those criteria may include the exceeding of speech envelope thresholds but also interposing a hangtime during which the currently active path may retain control, regardless of the activity in the other path. The inbound and outbound hangtimes may in embodiments be fixed or dynamic, and may be incremented or decremented depending on conditions. For instance, during periods of increasing noise or other parameters, either or both of the hangtimes may be incremented, or during periods of decreasing noise or other parameters, either or both of the hangtimes may be decremented. Greater continuity in speech or other interaction may therefore be achieved.

15

[00051]   Fig. 9(A) illustrates speech samples from speaker 120 and Fig. 9(B) illustrates speech samples from microphone 102 which may be processed in one regard according to the embodiment illustrated in Fig. 8. Fig. 10(A) illustrates the resulting outbound speech envelope (ob_env) along with the outbound break-in threshold (ob_break_in_thershold).

[00052]   Fig. 10(A) also illustrates the application of an outbound hangtime (ob_hangtime) interval during which the speaker path 130 may retain control and continue to be activated, despite the presence of energetic speech in the microphone path 128. Conversely, Fig. 10(B) illustrates the inbound speech envelope (ib_env) along with the inbound break-in threshold (ib_break_in_thershold). Fig. 10(B) also illustrates the application of an inbound hangtime (ib_hangtime) interval during which the microphone path 128 may retain control and continue to be activated, despite the presence of energetic speech in the speaker path 130. The introduction of these delay intervals may increase the sense of continuity for the near-end and far-end users during speakerphone operation.

[00053]   In particularly noisy environments, such as for example in urban areas, when an automobile window may be open, during playback of a noisy voice message or at other times, the fricatives and other signal components may tend to trigger the speaker path 130 to be muted, even when still-intelligible speech is present. This may in one regard be due to the crossing of an outbound muting threshold ordinarily intended to switch the speaker path 130 off when the far-end user input has degraded into noise. In an embodiment of the invention illustrated in Fig. 11, this effect may be addressed in one regard by eliminating the outbound off threshold (ob_off_threshold) and permitting the speaker path 130 to occupy the channel until the microphone path 128

16

contains energetic speech, rather than configuring the speaker path 130 to switch itself off below that threshold.

[00054]   As shown in that figure, processing may begin in step 1102. In step 1104, near-end samples from the microphone 102 may be processed by the speech encoder 108. In step 1106, outbound speech from the far-end user may be processed by speech decoder 124. In step 1108, the echo canceller 106 may receive the outputs of the speech encoder 108 and the speech decoder 124 to suppress echo and other feedback artifacts. In step 1110, the echo-cancelled inbound speech and the decoded outbound speech may be communicated to inbound speech envelope generator 132 and outbound speech envelope generator 134, respectively, to generate speech energy envelopes or other functions.

[00055]   In step 1112, an inbound on threshold (ib_on_threshold) and outbound on threshold (ob_on_threshold) may be generated, for instance similarly to the embodiment illustrated in Fig. 3 or otherwise. In step 1114, the duplex arbiter 1116 may apply control logic to lock to the microphone path 128 or the speaker path 130, according to the current speech envelopes of the paths.

[00056]   In step 1116, a determination may be made whether the outbound envelope (ob_env) exceeds the outbound on threshold (ob_on_threshold). If the outbound envelope (ob_env) does not exceed the outbound on threshold (ob_on_threshold), processing may proceed to step 1118 where a determination may be made whether the inbound envelope (ib_env) exceeds the inbound on threshold (ib_on_threshold). If the inbound envelope (ib_env) exceeds the inbound on threshold, processing may proceed to step 1120 where a determination may be made whether the speaker path 130 is locked, that is, currently has control of the communications channel, such as a

17

wireless cellular or other connection. If the speaker path 130 is locked, the state of the microphone path 128 and speaker path 130 may remain unchanged from the start of processing at step 1102 and control may proceed to step 1128 where processing for the current frame may end, following which processing may repeat, proceed to other tasks or end.

[00057]   If the determination at step 1120 is that the speaker path 130 is not locked, processing may proceed to step 1122 where the speaker path 130 may be deactivated or muted, while the microphone path 128 may be activated or unmuted. Processing then may likewise proceed to step 1128 to repeat, proceed to other tasks or end.

[00058]   If the determination at step 1118 is that the inbound envelope (ib_env) does not exceed the inbound on threshold (ib_on_threshold), processing may proceed to step 1128 to repeat, proceed to other tasks or end.

[00059]   If the determination at step 1116 is that the outbound envelope (ob_env) exceeds the outbound on threshold (ob_on_threshold), processing may proceed to step 1124 where a determination may be made whether the microphone path 128 is locked. If the microphone path 128 is not locked, control may proceed to step 1126 where the speaker path 130 may be activated or unmuted while the microphone path 128 may be deactivated or muted. Processing then may proceed to step 1128 to repeat, proceed to other tasks or end. Likewise, if the determination at step 1124 is that the microphone path 128 is locked, the state of the microphone path 128 and speaker path 130 may remain unchanged from the start of processing at step 1102, and control may proceed to step 1128 to repeat, proceed to other tasks or end.

[00060]   Fig. 12(A) illustrates samples from speaker 120 containing fricatives and other noise components, and Fig. 12(B) illustrates samples from microphone 102 at the

18

same time which may together be processed for instance according to the embodiment illustrated in Fig. 11. Fig. 13 illustrates speakerphone control which might occur when operating upon such signals without the benefit of the invention, including rapid switching of the speaker path 130 between on and off states, due to the fricative and other noise artifacts.

[00061]   Fig. 14(A) on the other hand illustrates the resulting speakerphone operation according to the embodiment of the invention illustrated in Fig. 11, in which the speaker path 130 maintains control of the channel even during relatively noisy background periods, in part because the outbound off threshold is eliminated, allowing the speaker path 130 to remain active. Instead of choppy or punctuated switching, the speaker path remains activated until the microphone path 128 appropriately seizes control of the channel due to energetic speech exceeding the inbound on threshold, as illustrated in Fig. 14(B). Smoother more continuous conversation results.

[00062]   The foregoing description of the system and method for speakerphone operation according to the invention is illustrative, and variations in configuration and implementation will occur to persons skilled in the art. For instance, while the invention has generally been described as containing discrete voice detectors in the form of inbound VAD 114 and outbound VAD 118, in embodiments the functions or parts of the functions of the two voice activity detectors could be combined in one part, or in one software module. More than two paths could also be managed according to the invention. Similarly, while the invention has been described with respect to an inbound path including an echo canceller 106, in embodiments other

19

types of noise suppressors could be implemented, or in embodiments that component could be omitted or modified.

[00063]   It has likewise been noted that the communications device in which the invention may operate may be or include a cellular telephone, but could consist of other communications platforms such as wired or wireless telephones, two-way radios, base stations for wireless telephones, network-enabled wireless communications devices such as 802.11a, 802.11b, 802.11g or other short or long-range telephony or other units, or other equipment as well.

[00064]   Yet further, while the invention has generally been described in terms of a speakerphone architecture in which the electronic intelligence governing the speakerphone operation is integral with the cellular telephone or other communications device, in other embodiments the intelligence may be embedded or shared in an attachment coupled to the communications device. For instance, the intelligence may be embedded or shared in a detachable battery, a headphone device, a tabletop or other fixed or non-wearable speakerphone unit, or in other accessories or parts. For example, the intelligence may enable a speakerphone operation through a car audio system coupled to a cellular telephone.

[00065]   In the case of a detachable or coupleable unit which adds or enhances speakerphone capability in a communications device, the intelligence embedded in the add-on device may communicate with the electronics of the communications device through interfaces such as a serial port such as an RS-232, a universal serial bus (USB) or a universal asynchronous receiver/transmitter (UART) connection, an infrared data (IrDA) port, a radio frequency link, or other serial, parallel or other data

20

ports or other connections. The scope of the invention is accordingly intended to be limited only by the following claims.